Known unknowns, unknown unknowns and information flow: new concepts and challenges in decentralized control

M.-A. Belabbas

***Abstract*— We introduce and analyze a model for decentralized control. The model is broad enough to include problems such as formation control, decentralization of the power grid and flocking. The objective of this paper is twofold. First, we show how the issue of decentralization goes beyond having agents know only part of the state of the system. In fact, we argue that a complete theory of decentralization should take into account the fact that agents can be made aware of only part of the global objective of the ensemble. A second contribution of this paper is the introduction of a rigorous definition of information flow for a decentralized system: we show how to attach to a general nonlinear decentralized system a unique *information flow graph* that is an invariant of the system. In order to address some finer issues in decentralized system, such as the existence of so-called ”information loops”, we further refine the information flow graph to a simplicial complex—more precisely, a Whitney complex. We illustrate the main results on a variety of examples.**

I. Introduction

Informally speaking, a decentralized control system is a system whose different parts—let us call the different parts agents—are not told what to do by a unique, centralized controller, but decide what to do based on the possibly *incomplete information* that is at their disposal.

The importance of decentralization in control has been recognized for many decades [1], but only in more recent time has the issue been the subject of sustained investigation, see [2], [3], [4], [5], [6], [7] and references therein. This renewed interest is fuelled, on the one hand, by the potential a complete theory of decentralized control has in explaining natural behavior: flocks of birds, ant colonies, etc. and more broadly by its applications in decision theory [8] and cognition. On the other hand, a theory of decentralization is also a necessity for engineering design: from smart grids to vehicles management on the highway [9], [10], recent developments in robotics and communication have made it possible to envision very large groups of autonomous vehicles or agents collaborating to achieve a global objective. In this context, decentralization is thus necessary for reasons ranging from robustness— failure at some level (agent, controller, etc.) in a centralized system is likely to affect the entire system, whereas failure in a decentralized system is more easily handled—to, at a more fundamental level, feasibility. Indeed, a centralized controller for, say, vehicles on the highway is not easily implemented.

The objective of this paper is twofold. First, we will show that the idea of *incomplete information* in a decentralized setting should be explored beyond the usual partial knowledge of the state of the system. While most of the extant work in decentralized control implicitly assume that all the agents know the objective of the ensemble and are thus solely constrained by their limited observations, we develop here a model which includes restrictions on what agents know about the global objective of the ensemble. We illustrate this idea on a formation control problem in [11] by showing that local stabilization around a given configuration is possible only if agents know more than their selfish objective.

Second, we will make rigorous the notion of *information flow in a decentralized system*. The naive notion of information flow that is often used in the linear theory of decentralized system—i.e. splitting variables into groups and coding the dependence between groups by a graph, see Section IV—is inherently dependent on the choice of coordinates used. This aspect puts it at odds with the idea that what an agent knows about the system should not depend on the way one chooses to describe the system. Even more, the naive information flow does not acknowledge the possibility that Lie brackets may be needed to make the system controllable. While these issues can be sidestepped in the linear case to obtain results that are nevertheless meaningful, they become a genuine limitation when one tries to understand nonlinear decentralized systems or systems with constraints. Indeed, in the former situation, there often does not exist preferred coordinates or one may need several coordinate charts to describe the system. In the latter situation, choosing coordinates that are compatible with the constraints, e.g. a conservation of energy constraint, changes the naive information flow since it often requires a mixing of the coordinates used to describe the agents individually.

We introduce in this paper a definition of information flow graph that is invariant under changes of coordinates and allows to define rigorously decentralization in a control system. The main idea behind this definition is that the observation functions on the system provide a natural set of vertices for the information flow graph.

The paper is organized as follows. We first introduce a general nonlinear model for decentralized control systems. We then define the *global objective* of a system and the *local or selfish objectives* of the agents. The fact that the agents only know part of the global objective is enforced through the use of non-invertible functions on the parameters describing the global objective; this approach can be understood, informally speaking, as a *decentralization of the objective* or a decentralization of the design of the control law.

M.-A. Belabbas is with the School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138 belabbas@seas.harvard.edu

In the following section, we introduce a partial order on the local objectives and observation functions. This partial order allows us to quantify the idea that some local objectives (resp. observations) are more revealing of the global objective (resp. state of the ensemble) than others.

In Section IV, we introduce a coordinate free definition of information flow in a decentralized system. Motivated by the existence of ”information loops” in decentralized system, we further refine the information flow graph into an information flow complex that reveals finer issues in decentralization.

Example 1 (Power grid)**.** *In recent years, the development of methods to insure the stability of the power grid have come at the forefront of research in control theory. In this context, one can view the power grid as a very large scale system whose global objective is to remain stable around a desired operating point. Such systems currently operate in a centralized manner under the supervision of Supervisory Control and Data Acquisition system (SCADA). This centralized framework, however, is starting to show its limitations due to the increasingly complex components that are part of the grid (e.g. green energy suppliers).*

The development of decentralized methods to insure the stability and good operation of the grid have thus become a priority in power systems engineering. In this context, the global objective is a function of all the components of the grid, but it is clearly not feasible to let all agents in the grid know about its complete architecture.

Example 2 (Formation control)**.** *Let $x_i \in \mathbb{R}^2$ represent the positions of autonomous agents in the plane and $d_i \in (0,\infty)$ be real positive constants.*

We consider the formation control problem whose dynamics are given by

$$\begin{aligned}\dot{x}_1 &= e_1(x_2-x_1)+e_5(x_4-x_1)\\ \dot{x}_2 &= e_2(x_3-x_2)\\ \dot{x}_3 &= e_3(x_1-x_3)\\ \dot{x}_4 &= e_4(x_3-x_4)\end{aligned}$$

where we denote by e_i the error in edge length:

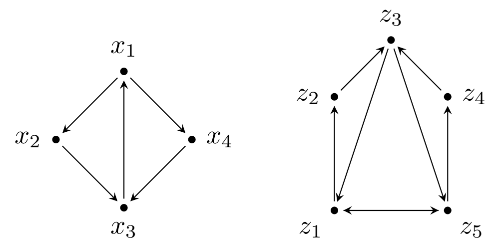
$$e_1 = \|x_2-x_1\|^2 - d_1, e_2 = \|x_3-x_2\|^2-d_2,\ldots,\\ e_4 = \|x_3-x_4\|^2-d_4, e_5 = \|x_4-x_1\|^2-d_5$$

The information flow of the system is represented in Figure 1a. Formation control problems are defined up to a rigid transformation of the plane [12]. For this reason, one often describes the dynamics in terms of the inter-agent distances [13]

$$\left\{\begin{aligned} z_1 &= x_2-x_1\\ z_2 &= x_3-x_2\\ z_3 &= x_1-x_3\\ z_4 &= x_3-x_4\\ z_5 &= x_4-x_1,\end{aligned}\right. \tag{1}$$

instead of the absolute positions of the agents.

In the z variables, the dynamics of the system are given by



(a) Naive information flow for the x variables

(b) Naive information flow for the z variables

Fig. 1: The naive information flow of a system is a directed graph whose vertices v_i correspond to groups of variables describing the system. There is a directed edge from v_i to v_j if the dynamics of variables in the group of v_i depend on the variables in the group of v_j. This graph depends on the coordinates chosen to describe the system.

$$\begin{aligned}\dot{z}_1 &= e_2z_2-e_1z_1-e_5z_5\\ \dot{z}_2 &= e_3z_3-e_2z_2\\ \dot{z}_3 &= e_1z_1+e_5z_5-e_3z_3\\ \dot{z}_4 &= e_3z_3-e_4z_4\\ \dot{z}_5 &= e_4z_4-e_1z_1-e_5z_5\end{aligned}$$

The corresponding naive information flow has 5 vertices and is represented in Figure 1b. We describe in Section IV a way to obtain the information flow depicted in Figure 1a from the description of the system given in terms of the z variables.

II. A MODEL FOR DECENTRALIZED CONTROL

We present in this section a general model for nonlinear decentralized control systems. The model is a natural extension of the notion of decentralized system that is encountered in the literature on linear systems.

In addition to being applicable to nonlinear problems, such as formation control, our approach distinguishes itself from most of the work on linear decentralized control in at least two major aspects [2], [3], [5]:

- it introduces the notion of parametrized objective of a decentralized system.
- it allows for *loops of information* in the system, unlike approaches based on quadratic invariance or partial orders [2], [3]. We revisit this point in Section IV.

A. General model

Let M be a smooth manifold and the state $x \in M$. We consider nonlinear control systems of the type

$$\dot{x} = f(x,u(x)) = \sum_{i=1}^{n} u_i(\delta_i(\mu); h_i(x))g_i(x) \tag{2}$$

where δ, h are smooth functions, the g_i's are smooth vector fields and μ is a parameter that describes the objective of the

system. We let $\mathcal{U}$ be the space of admissible controls u_i. We elaborate on the various parts of the model in this section.

A common situation is for the manifold M to be the product of the manifolds describing the state-spaces of each agent:

$$M = \bigotimes_{i=1}^{n} M_i$$

where M_i is the state-space of agent i. We can thus write that the tangent space of M is the direct sum $TM = \oplus_i TM_i$. In this case, we also have that the projection of $g_i(x)$ onto TM_j is zero if $i \neq j$:

$$\pi_j g_i(x) = 0 \text{ if } i \neq j.$$

This product structure is often lost due to either interactions between agents which impose constraints on the state $x \in M$, or the existence of a symmetry group acting on M, in which case one has to consider equivalence classes of states $x \in M$ (this is the case in, e.g., formation control). Hence we do not assume here any special structure for M.

We differentiate between two type of objectives:

1) the objective that each agent or plant tries to satisfy: it is referred to as local objective or selfish objective.
2) the objective the agents try to achieve by cooperating: it is referred to as global objective or common objective.

The functions δ_i in Equation (2) allow us to control how much an agent knows about the common objective of the ensemble; in some sense, these functions introduce a partial observation on the objective of the ensemble, akin to the partial observation that the agents have on the state of the ensemble. They are described in more detail below; we start with the definition of local observations.

B. Local observations

The main characteristic of a decentralized control system is that the agents are only able to observe part of the state of the system. We introduce the functions

$$h_i(x) : M \to \mathbb{R}^{k_i}, k_i \text{ a positive integer}$$

to describe the observation of agent i on the current state of the system. We denote by $h_i(M)$ the image of M under the map h_i.

C. Local and global objectives

We define in this section the *local and global objective of a decentralized system*. We consider the case of objectives that depend on a parameter. This level of generality is often necessary to accurately model decentralized systems whose dynamics are rich enough to accommodate parameter-varying objectives, such as flocks of autonomous agents or power distribution systems.

1) Global objective: Let P be a smooth manifold, we let $\mu \in P$ parametrize the global objective of the control system as follows:

Definition 1 (Global objective)**.** *Given a decentralized control system $\dot{x} = f(x, u(x))$ of the type of Equation* (2)*, the global objective function is a differentiable function*

$$F(\mu; x, u) : P \times M \times \mathcal{U} \to \mathbb{R}^d$$

with the convention that the objective is achieved if the system is at $x^* \in M$ with

$$F(\mu; x^*, u) = 0$$

for equality objectives *or*

$$F(\mu; x^*, u) \geq 0$$

for inequality objectives*, where the inequality is taken entry-wise.*

The objective function can in general depend on u; this dependence is necessary if ones considers stabilization objectives. When a global objective is not parametric or does not depend on u explicitly, we omit the dependence from the notation. We give a few examples:

Example 3 (Rendez-vous)**.** *Consider a multi-agent system with two agents whose positions are given by $x_1 \in \mathbb{R}^m$ and $x_2 \in \mathbb{R}^m$. The global objective is to have the agents meet. This objective does not depend on a parameter. We can encode it by*

$$F(x_1, x_2) = -\|x_1 - x_2\|^2.$$

If we want the agents to reach a position such that they are at a given distance d from each other, we let $P = [0, \infty)$ and we use

$$F(d; x_1, x_2) = -(\|x_1 - x_2\|^2 - d^2)^2.$$

Example 4 (Stabilization)**.** *Consider the simple nonparametric rendez-vous problem described in the previous example with the addition that the agents are required to stabilize at the rendez-vous configuration. We denote by*

$$\frac{\partial f}{\partial x}|_{x^*}$$

the Jacobian of the system at x^*. We denote by $\lambda_i(A)$ the eigenvalue of A with i^{th} largest real part. We can represent this global objective by using the vector-valued function

$$F(d; x, u) : \mathbb{R}^m \times \mathbb{R}^m \times \mathcal{U} \to \mathbb{R}^{m+1} : \\ x \to \begin{bmatrix} -(\|x_1 - x_2\|^2 - d^2)^2 \\ -Re(\lambda_1(\frac{\partial f}{\partial x})) \\ \vdots \\ -Re(\lambda_n(\frac{\partial f}{\partial x})) \end{bmatrix} \quad (3)$$

2) Local Objectives: We now focus on describing the system at the level of the agents. We define a *local objective* as being, roughly speaking, a restriction of a global objective.

Definition 2 (Local objective)**.** *Given a decentralized control system with global objective parametrized by P, we let P_i be a smooth manifold and*

$$\delta_i : P \to P_i$$

be smooth functions. For $\mu \in P$, the local objective of agent

i *is given by a smooth function*

$$F_i(\delta_i(\mu); h_i(x), u_i) : P_i \times h_i(M) \times \mathcal{U}_i \to \mathbb{R}^d$$

with the same convention as in Definition 1 regarding equality and inequality objectives.

When δ_i is not an invertible function, an agent *knows about part of the global objective* . We give some examples of relations between global and local objectives in the section below.

The decentralized control problem is *well-posed* if satisfying the local objectives is sufficient to satisfy the global objective:

$$f_i(\delta_i(\mu), h_i(x), u_i(x)) \geq 0 \text{ for all } i \implies F(\mu, x, f(x)) \geq 0$$

and similarly for the equality objective.

A wide array of questions in decentralized control can then be reduced to one on the following three major questions:

1) How little information can we let the agents know about the global objective and still have the ensemble achieve it? In other words, how informative do we need the δ_i to be in order to achieve a given global objective?
2) Given a global objective, how little observation on the system do the agents need in order to achieve the global objective? In other words, how informative do the h_i need to be in order to achieve a global objective?
3) Given a decentralized system with fixed observation functions $h_i(x)$ and control vector fields $g_i(x)$, what global objectives are achievable?

The first two questions are not independent. Indeed, if the observation functions $h_i(x)$ do not provide much information about the state of the ensemble, increasing the knowledge an agent has about the global objective is likely to be fruitless (a typical example is formation control). We introduce below a partial order on observations and objective that allow us to attach a mathematically precise meaning to these questions.

Example 5 (Formations)**.** *Consider a formation control problem where* n *agents in the plane, with positions* $x_i \in \mathbb{R}^2$*, are required to stabilize at the configuration described in Figure 2. This configuration is defined up to a translation and rotation of the plane. We have shown in [12] that the space of such configurations was* $\mathbb{C}\mathrm{P}(n-2)\times(0,\infty)$*. Hence, the parameter space is* $P = \mathbb{C}\mathrm{P}(n-2) \times (0, \infty)$*. A point in* P *can also be represented, up to mirror symmetry, by the inter-agent distances* $[d_1, \dots, d_N]$*, where* $N = \frac{1}{2}n(n-1)$*. The* d_i*'s are of course redundant in this representation, since simple trigonometric rules relate the pairwise distances. The problem of finding a non-redundant representation based on fewer pairwise distances is related to global rigidity [14]*

It may be impractical, or in some situation undesirable, to let every agent know about the complete vector μ*. The* δ_i *introduced here allow the study of systems where the amount of knowledge an agent gets about* μ *is controlled.*

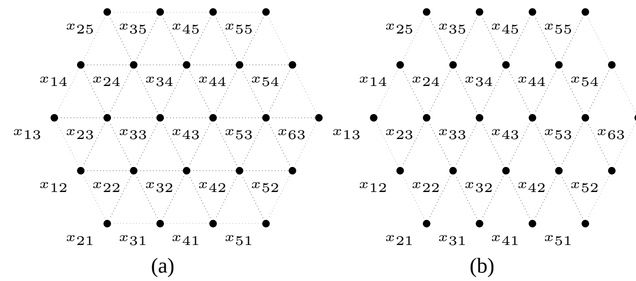


Fig. 2: Consider the formation control problem where agents with positions $x_{ij} \in \mathbb{R}^2$ are required to stabilize at the configuration depicted above. We let μ be the vector of all pairwise distances between agents at the desired configuration. We represent the function $\delta_i(\mu)$ by a graph with a vertex per agent, and an edge between the vertices x_{ij} and x_{kl} if agents x_{ij} and x_{kl} know the distance $\|x_{ij} - x_{kl}\|$ at the desired configuration. In order to have the agents cooperate to stabilize at this configuration, we can let each agent know about the complete vector μ or only parts of it. For example, letting each agent know about the distance to its nearest neighbors, as shown above in (a), allows them to reconstruct the desired configuration. In (b), the agents are given less information about the global objective than in (a), and one can see that respecting the pairwise distances depicted is not sufficient to reconstruct the desired configuration. Hence the δ_i in (b) are not informative enough.

III. Partial orders and decentralization

In the design of a decentralized control system, a global objective can be achieved by the use of different observation functions and different local objectives. We put a partial order on the local objectives (resp. observations) to formalize the notion that different local objectives (resp. observations) can be more or less revealing of the global objective or system (resp. state of the ensemble).

We start with the definition of partial order:

Definition 3 (Partial order)**.** *A partial order* $\succeq$ *over a set* F *is a binary relation on the elements of* F *which satisfies, for* $f_1, f_2, f_3 \in F$

1) *reflexivity:* $f_1 \succeq f_1$.
2) *antisymmetry: if* $f_1 \succeq f_2$ *and* $f_2 \succeq f_1$ *then* $f_1 = f_2$
3) *transitivity:* $f_1 \succeq f_2$ *and* $f_2 \succeq f_3$ *then* $f_1 \succeq f_3$

The elements f_1 and f_2 in F are called comparable if either $f_1 \succeq f_2$ or $f_2 \succeq f_1$ hold. The set F is called a *poset* for *partially ordered set*. An element $f \in F$ is a greatest (resp. smallest) element if $f \succeq f_i$ (resp. $f_i \succeq f$) for all $f_i \in F$. If a greatest (resp. smallest) element exists, it is unique.

A maximal (resp. minimal) element f is such that there is no $f_i \in F$ such that $f_i \succeq f$ (resp. $f \succeq f_i$).

Remark 1. *We mention here that partial orders have, quite interestingly, been applied to decentralized control in previous work [3]. However, the approach and objective are quite*

different from ours. In the work [3], the authors give an analysis of linear systems whose information flow graph—we will define it in Section IV—is given by a Hasse diagram [15]. These are a type of directed acyclic *graphs. They show in particular, relying on [2], that there is a parametrization of such systems in which stabilization questions can be reduced to convex problems.*

A. Partial order on δ_i

Recall that the functions δ_i allow us to define decentralized systems where the agents know only part of the global objective of the ensemble. We further refine this notion of incomplete knowledge of the global objective by establishing a *partial order* of the functions δ_i. Let $\mu \in P$, where P is a smooth compact manifold. From Whitney's embedding theorem [16], we know that for n large enough, we can smoothly embed P in $\mathbb{R}^n$. Hence, without loss of generality, we can assume that all the δ_i map into $\mathbb{R}^n$. We denote by $N_\delta(\mu)$ the isolevel set

$$N_\delta(\mu) = \{x \in P \text{ s.t. } \delta(x) = \delta(\mu)\}.$$

Many functions δ_i describe a similar restriction of the objective. For example, if δ_i maps to $\mathbb{R}$, translating the function by a constant $c \in \mathbb{R}$ to $\delta_i(\mu) + c$ does not change, for all practical purposes, what agent i knows about the global objective. Indeed, if the objective is realized with the control u_i for δ_i, it is realized with the control $\tilde{u}_i(\delta_i; x) = u_i(\delta_i - c; x)$ for $\delta_i + c$. We generalize this idea in the following definition:

Definition 4 (Equivalence of local objectives)**.** *The functions $\delta_1 : P \to \mathbb{R}^n$ and $\delta_2 : P \to \mathbb{R}^n$ are equivalent at $\mu \in P$, written $\delta_1 \approx_\mu \delta_2$, if*

$$N_{\delta_1}(\mu) = N_{\delta_2}(\mu).$$

They are equivalent if the above is true for all $\mu \in P$.

Hence two functions are equivalent if their isolevel sets are the same. In particular, all one-to-one invertible functions are equivalent. This definition indeed corresponds to the intuitive notion of equivalent knowledge of the global objective: as explained above, if the isolevel sets of δ_1 and δ_2 are the same, one can realize the same decentralized system by using an appropriately modified control u.

Example 6. *Assume that we have a 2 agent decentralized system in $\mathbb{R}^m$ where the objective is parametrized by a point in $P = [0,1] \times [0,1]$. Let $\mu = (\mu_1, \mu_2) \in P$ and $\delta_1(\mu) = \mu_1$, $\delta_2(\mu) = \mu_2$. The uncertainty the first agent has about the global objective is its uncertainty about μ_2. In particular, if agent 1 has access to $\tilde{\delta}_1(\mu) = \mu_1^2$, it has the same knowledge about the global objective than with $\delta_1(\mu)$. According to Definition 6, $\delta_1 \approx \tilde{\delta}_1$. Observe that this would not be true if $P = [-1,1] \times [-1,1]$.*

We can now define a partial order on the local objectives.

Definition 5 (Partial order on δ_i.)**.** *Let F be the set of continuous functions on P with the equivalence relation of Definition 4. We say that*

$$\delta_1 \succeq_\mu \delta_2 \textit{ if } N_{\delta_1}(\mu) \subseteq N_{\delta_2(\mu)}.$$

If the above relation is valid for all $\mu \in P$, we simply write

$$\delta_1 \succeq \delta_2.$$

This definition expresses the notion that δ_1 contains more information than δ_2—written as $\delta_1 \succeq \delta_2$— if the uncertainty arising from knowing $\delta_1(\mu)$ is smaller than the one arising from knowing $\delta_2(\mu)$, where uncertainty is quantified by the isolevel sets of δ.

The partial order $\succeq$ has a smallest element: the constant function. This corresponds to the intuitive idea that if δ_i is constant, the agent knows nothing about the global objective. The functions with the highest level of information are the invertible functions of μ; these functions are the maximal elements.

Example 7. *Assume that*

$$P = S^3 = \{x \in \mathbb{R}^4 \textit{ s.t. } x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1\}.$$

We let $\delta_1(x) = x_1$ and $\delta_2(x) = (x_1, x_2)$.We have that

$$\delta_1 \succeq \delta_2.$$

We also have

$$\delta_2 \approx (x_2, x_1) \approx (x_1 + a, x_2 + b),$$

for $a, b \in \mathbb{R}$

B. Partial order on h_i

We similarly define a partial order on the observations h_i. We denote by $N_{h_i}(x)$ the subset of M such that $h_i(y) = h_i(x)$:

$$N_h(x) = \{y \in M \text{ s.t. } h(y) = h(x)\}$$

In words, it is the set of configurations that are *undistinguishable* to the observation function h_i.

Similarly to Definition 4, we say that two observation functions h_1 and h_2 are equivalent if their isolevel sets on M—or the configurations that are undistinguishable for h_1 and h_2—are the same:

$$h_1 \approx h_2 \Leftrightarrow N_{h_1}(x) = N_{h_2}(x), \forall x \in M.$$

Furthermore, we can use a similar partial ordering on the observation functions to the one of Definition 5:

$$h_1 \succeq_x h_2 \text{ if } N_{h_1}(x) \subseteq N_{h_2(x)}.$$

If the above relation is valid for all $x \in M$, we simply write

$$h_1 \succeq h_2.$$

C. Partial order on the f_i

We further define a partial order on the f_i. In the case of equality objectives, the defintions are similar to the ones we have introduces above. We thus treat here the case of inequality objectives. We denote by $N^+_{f_i(\delta_i; h_i, u_i)}(\mu)$ the

subset of M_i such that $f_i(\delta_i(\mu); h_i(x), u_i(x)) \geq 0$:

$$N^+_{f_i(\delta_i;h_i,u_i)}(\mu) = \{x \in M \text{ s.t. } f_i(\delta_i(\mu); h_i(x), u_i(x)) \geq 0\}$$

In words, it is the set of configurations that satisfy the local objective f_i.

We thus introduce the equivalence relation

Definition 6. *The functions f_1 and f_2 are equivalent at $\mu \in P$ if*

$$f_1 \approx^+_\mu f_2 \Leftrightarrow N^+_{f_1(\delta_i,h_i,u_i)}(\mu) = N^+_{f_2(\delta_i,h_i,u_i)}(\mu).$$

They are equivalent if the above is true for all $\mu \in P$.

We say that

$$f_i \succeq_\mu f_j \text{ if } N^+_{f_i(\delta_i,h_i,u_i)}(\mu) \subseteq N^+_{f_j(\delta_i,h_i,u_i)}(\mu).$$

Hence, $f_1 \succeq f_2$ if the local objective f_1 is *more stringent* than the local objective f_2.

D. Minimally informed decentralized control and saturation

We can now define

Definition 7 (Minimally informed decentralized system)**.** *Given a global objective F, we say that a decentralized control system of the type of Equation* (2) *is minimally informed if there is no set of functions $\tilde{\delta}_i, \tilde{h}_i, \tilde{f}_i$ with*

$$\delta_i \succeq \tilde{\delta}_i, h_i \succeq \tilde{h}_i, f_i \succeq \tilde{f}_i$$

and such that the decentralized system

$$\dot{x} = \sum_i u_i(\tilde{\delta}_i(\mu), \tilde{h}_i(x)) g_i(x)$$

with local objectives $\tilde{f}_i$ satisfies the global objective F.

When looking for a minimally informed system, an important notion that arises is the one of saturation. Since the agents only have access to partial observations on the system, knowing an increasingly larger part of μ may cease to be helpful.

We say that δ_i saturates h_i if for all $\tilde{\delta}_i \succeq \delta_i$, there are no $\tilde{f}_i(\tilde{\delta}_i(\mu); h_i(x))$ with

$$\tilde{f}_i \succeq f_i.$$

Reciprocally, we have that h_i saturates δ_i if for all $\tilde{h}_i \succeq h_i$, there are no $\tilde{f}_i(\delta_i(\mu); \tilde{h}_i(x))$ with $\tilde{f}_i \succeq f_i$.

Example 8. *Consider agent 1 in the two-cycles formation of Figure 3. We let μ denote a target formation as explained in Example 5. We let $\delta_1(\mu) = [d_1, d_5]$. In the case of range only measurements*

$$h_1(x) = [\|x_2 - x_1\|, \|x_4 - x_1\|],$$

and the local objective is given by

$$f_1(\delta_1(\mu); h_1(x)) = \begin{bmatrix} \|x_2 - x_1\| - d_1 \\ \|x_4 - x_1\| - d_5 \end{bmatrix}.$$

It is easy to see that h_1 is saturated by δ_1. Similarly, for agent 2 *with*

$$h_2(x) = [\|x_2 - x_1\|],$$

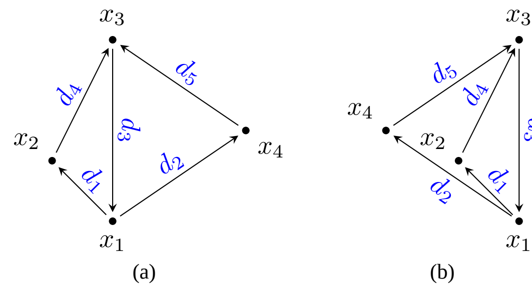


Fig. 3: In the two-cycles formation depicted above, the agents are required to stabilize at the inter-agent distances $d_1, \ldots, d_5$. Up to mirror symmetry, there are two configurations in the plane that satisfy these interagent distances. If agent 1 can measure the *relative positions* of agents 2 and 4, it can make use of the angle between the vectors $(x_2 - x_1)$ and $(x_4 - x_1)$ at the desired configurations (a) and (b). This observation function is thus not saturated by d_1, d_5. If agent 1 can only measure its distance to agent 1 and agent 2, the knowledge of the angle is not helpful. This observation function is saturated by d_1, d_5.

$\delta_2(\mu) = [d_2]$ and $f_2(\delta_2(\mu); h_2(x)) = \|x_2 - x_1\| - d_1$, we see that $\delta_2(\mu) = d_2$ saturates h_2.

If we let

$$h_1(x) = \left[\|x_2 - x_1\|, \|x_4 - x_1\|, (x_2 - x_1)^T(x_4 - x_1)\right],$$

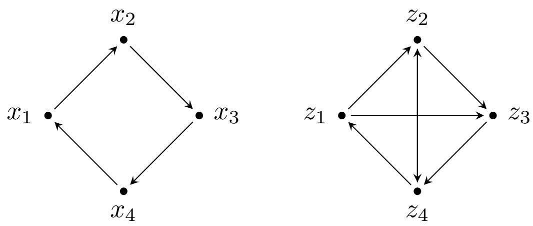
i.e. the first agent observes its relative distances to agents 1 and 4 as well as their relative positions, then additional knowledge of μ is helpful. Indeed, we can prove in this case that h_1 is saturated by $\delta_1(\mu) = \mu$. Intuitively, knowing all these distances allows agent 1 *to establish what the possible angles between $x_2 - x_1$ and $x_4 - x_1$ are when the global objective is reached. See [11] for additional details.*

IV. Information flow graph

The abstract idea of information flow, because it allows to grasp the connectivity of the agents and identify potential difficulties in the distribution of information in the ensemble, appears quite frequently in work on decentralized control.

We now define what we informally call the *naive information flow* of a system; which is oftentimes implicitly defined in work on decentralized control. Let e_j be the vector with zero entries except for the j^{th} entry, which is one. The e_j's form the canonical basis of $\mathbb{R}^n$. In the case of multi-agent systems in $\mathbb{R}^n$, $h_i(x)$ will often be the projection of $x \in \mathbb{R}^n$ onto a the subspace spanned by some vectors e_j, $j \in \mathcal{J}_i$ where $\mathcal{J}_i$ is a set of indices. For this reason, the observation functions h_i are encoded as a graph with vertices x_i and an edge from x_i to x_l if $l \in \mathcal{J}_i$. We call this the *naive* information flow of the system, since it is coordinate dependent as we illustrate in the examples below.

Example 9 (Four agents)**.** *Consider the system with $M = \mathbb{R}^{4m}$, $x_i \in \mathbb{R}^m$ and whose dynamics is given by*



(a) (b)

Fig. 4: The two graphs above represent the naive information flow of the same system expressed in two different coordinate systems.

$$\begin{aligned}
\dot{x}_1 &= u_1(x_1, x_2) \\
\dot{x}_2 &= u_2(x_2, x_3) \\
\dot{x}_3 &= u_3(x_3, x_4) \\
\dot{x}_4 &= u_4(x_4, x_1)
\end{aligned}$$

Hence, $\mathcal{J}_1 = \{1,2\}, \mathcal{J}_2 = \{2,3\}, \mathcal{J}_3 = \{3,4\}, \mathcal{J}_4 = \{1,4\}$. *One can associate the graph of Figure 4a to this system, which shows a non-trivial loop in the information flow. Now consider the linear change of variables:*

$$\begin{aligned}
z_1 &= x_1 + x_2 + x_3 + x_4 \\
z_2 &= x_2 + x_3 + x_3 \\
z_3 &= x_3 + x_4 \\
z_4 &= x_4
\end{aligned}$$

We then have

$$\begin{aligned}
x_1 &= z_1 - z_2 \\
x_2 &= z_2 - z_3 \\
x_3 &= z_3 - z_4 \\
x_4 &= z_4
\end{aligned}$$

and for appropriately defined $\tilde{u}_i$,

$$\begin{aligned}
\dot{z}_1 &= \tilde{u}_1(z_1, z_2, z_3) \\
\dot{z}_2 &= \tilde{u}_2(z_2, z_3, z_4) \\
\dot{z}_3 &= \tilde{u}_3(z_3, z_4) \\
\dot{z}_4 &= \tilde{u}_4(z_1, z_2, z_4).
\end{aligned}$$

In this case, $\mathcal{J}_1 = \{1,2,3\}, \mathcal{J}_2 = \{2,3,4\}, \mathcal{J}_3 = \{3,4\}, \mathcal{J}_4 = \{1,2,4\}$. *This system corresponds to the naive information flow graph depicted in Figure 4b.*

A. Information flow graph of a decentralized system

We have seen above that the naive definition of information flow is not satisfactory since it depends on the parametrization chosen for the system, whereas decentralization is a coordinate-free notion: our choice of coordinates to describe a system should not affect the knowledge each agent has about the system.

We provide here a coordinate free definition of information flow. The idea is to let the observation functions h_i, define the vertices of the graph, and use the vector fields g_i to determine the presence of edges, Precisely, to a decentralized control system of the type

$$\dot{x} = \sum_i u_i(\delta_i(\mu); h_i(x)) g_i(x)$$

we will assign a directed graph at first, and refine the notion to obtain a *simplicial complex*.

Recall that a vector field $g(x)$ on M acts on functions h defined on M via differentiation. We write this action as

$$g \cdot h(x).$$

If $h(x) = [h^1, \ldots, h^k]$ is vector-valued, we define $g \cdot h$ as $g \cdot h = [g \cdot h^1, \ldots g \cdot h^k]$.

For example, on $\mathbb{R}^n$ with coordinates $(x_1, \ldots, x_n)$, the vector fields $g_1(x) = [g_{11}(x), \ldots, g_{1n}(x]$ and $g_2(x) = [g_{21}(x), \ldots, g_{2n}(x)]$ act on the function $h(x)$ via

$$g_i(x) \cdot h = \sum_{j=1}^{n} g_{ij} \frac{\partial}{\partial x_j} h.$$

The Lie bracket of g_1 and g_2 is the vector field

$$[g_1, g_2](x) = \frac{\partial g_2}{\partial x} g_1 - \frac{\partial g_1}{\partial x} g_2$$

where $\frac{\partial g}{\partial x}$ is the Jacobian matrix of g.

Definition 8 (Information Flow Graph)**.** *Consider the decentralized control system*

$$\dot{x} = \sum_{i=1}^{n} \sum_{j=1}^{n_i} u_{ij}(\delta_i(\mu); h_i(x)) g_{ij}(x) \tag{4}$$

where all the functions and vector fields involved are smooth. We assign to this system the graph with n *vertices* $h_1, h_2, \ldots, h_n$ *and edges given according to the following rules:*

$n_i = 1$ *there is an edge from* h_j *to* h_i *if*

$$g_{ik}(x) \cdot h_j(x) \neq 0 \text{ for any } k = 1 \ldots n_i$$

$n_i \neq 1$ *Let* $\{g_{i1}, \ldots, g_{in_i}\}_{LA}$ *be the set of vector fields obtained by taking iterated Lie brackets of* $g_{i1}, \ldots, g_{in_j}$. *There is an edge between from* h_j *to* h_i *if*

$$g_{ik}(x) \cdot h_j(x) \neq 0 \text{ for any } g_{ik} \in \{g_{i1}, \ldots, g_{in_i}\}_{LA}.$$

In words, there is an edge from h_j *to* h_i *if the motion of an agent that uses the observation function* h_i *is observable by* h_j.

In the multi-agent case, each agent will often have its own observation function, and the above can be rephrased as saying that there is an edge from agent i to agent j if agent i can observe changes in the state of agent j.

Example 10. *Consider the system on $\mathbb{R}^4$ given by*

$$\begin{aligned}\dot{x}_1 &= u_1(x_1, x_2)\\ \dot{x}_2 &= u_2(x_2, x_3)\\ \dot{x}_3 &= u_3(x_3, x_4)\\ \dot{x}_4 &= u_4(x_4, x_1)\end{aligned}$$

We let $x = (x_1, x_2, x_3, x_4)$ and

$$h_1(x) = (x_1, x_2), h_2(x) = (x_2, x_3), h_3(x) = (x_3, x_4) \text{ and } h_4(x) = (x_1, x_4).$$

We define

$$g_1(x) = [1, 0, 0, 0], g_2(x) = [0, 1, 0, 0], \ldots, g_4(x) = [0, 0, 0, 1]$$

The system of Equation (5) *can thus be written as*

$$\dot{x} = \sum_i u_i(h_i(x))g_i(x).$$

According to definition 12, the information flow graph is given by $G = (V, E)$ with

$$V = \{h_1, h_2, h_3, h_4\}$$

and

$$E = \{(h_1, h_2), (h_2, h_3), (h_3, h_4), (h_4, h_1)\}.$$

The same system expressed in the z variables defined in Example 9 is given by

$$\begin{aligned}\dot{z}_1 &= u_1(z_1 - z_2, z_2 - z_3) + u_2(z_2 - z_3, z_3 - z_4)\\ &\quad +u_3(z_3 - z_4, z_4) + u_4(z_4, z_1 - z_2)\\ \dot{z}_2 &= u_2(z_2 - z_3, z_3 - z_4) + u_3(z_3 - z_4, z_4)\\ &\quad +u_4(z_4, z_1 - z_2)\\ \dot{z}_3 &= u_3(z_3 - z_4, z_4) + u_4(z_4, z_1 - z_2)\\ \dot{z}_4 &= u_4(z_4, z_1 - z_2).\end{aligned}$$

The observation functions are

$$\tilde{h}_1(z) = (z_1 - z_2, z_2 - z_3), \tilde{h}_2(z) = (z_2 - z_3, z_3 - z_4), \tilde{h}_3(z) = (z_3 - z_4, z_4), \tilde{h}_4(z) = (z_4, z_1 - z_2).$$

The control vector fields become

$$\tilde{g}_1(z) = [1, 0, 0, 0], \tilde{g}_2(z) = [1, 1, 0, 0], \tilde{g}_3(z) = [1, 1, 1, 0] \text{ and } \tilde{g}_4(z) = [1, 1, 1, 1].$$

We can now write

$$\dot{z} = \sum_i u_i(\tilde{h}_i(z))\tilde{g}_i(z).$$

The information flow graph associated to this system has four vertices $\tilde{h}_1, \tilde{h}_2, \tilde{h}_3, \tilde{h}_4$ We have the following relations

$$\begin{aligned}\tilde{g}_1 \cdot \tilde{h}_2 &= [\partial_{z_1}(z_2 - z_3), \partial_{z_1}(z_3 - z_4) = [0, 0]\\ \tilde{g}_1 \cdot \tilde{h}_3 &= [0, 0];\ \tilde{g}_1 \cdot \tilde{h}_4 = [0, 1]\\ \tilde{g}_2 \cdot \tilde{h}_1 &= [0, 1];\ \tilde{g}_2 \cdot \tilde{h}_3 = [0, 0]\\ \tilde{g}_2 \cdot \tilde{h}_4 &= [0, 0];\ \tilde{g}_3 \cdot \tilde{h}_1 = [0, 0]\\ \tilde{g}_3 \cdot \tilde{h}_2 &= [0, 1];\ \tilde{g}_3 \cdot \tilde{h}_4 = [0, 0]\\ \tilde{g}_4 \cdot \tilde{h}_1 &= [0, 0];\ \tilde{g}_4 \cdot \tilde{h}_2 = [0, 0]\\ \tilde{g}_4 \cdot \tilde{h}_3 &= [0, 1]\end{aligned}$$

These relations yield the same information flow graph as above.

B. Decentralized systems and Whitney complex

Definition 12 in the previous section attaches an information flow graph to a decentralized system in a coordinate free manner. The salient point was that the observation functions $h_i(x)$ provide a natural set of vertices for the graph.

Consider the triangular formation of Figure 5a. The main source of difficulty in the control of this formation comes from the fact that the motion of x_i depends on the motion of x_{i+1} (taken modulo 3): if x_2 moves, x_1 has to adjust itself, which forces x_3 to move which in turn provokes a motion of x_2. We call this a nontrivial *loop of information*.

Now consider the triangular formation with bidirectional edges. The above mentioned loop of information still exists, but its effect on the dynamics is diluted due to the fact that the communication goes both ways between the agents. In fact, there is all-to-all communication between agents in this formation, and the system is thus equivalent to a centralized one, where each agent implements locally a copy of a centralized controller. We call this information loop trivial. We devote the remainder of this section to putting this notion of triviality on a firm mathematical footing.

In order to do so, we need some concepts from algebraic topology. The role of homological algebra and algebraic topology in control theory and applied sciences has been recognized in many different contexts such as feedback stabilization, computer graphics, sensor networks or data analysis [17], [18], [19], [20]. We show here how related ideas naturally appear in the definition of decentralized systems.

We start with some graph theoretic definitions. We recall here that the information flow graph is in general a mixed graph (i.e. containing both directed and undirected edges). We say that $G = (V, E)$ is an *undirected complete graph* if

$$E = \{(v_i, v_j) \text{ s.t. } v_i, v_j \in V\}.$$

In words, G contains all possible undirected edges on its vertices. If $G = (V, E)$ is a graph, we call $G' = (V', E')$ the subgraph of G generated by $V' \subset V$ when $E' \subset E$ is the set of edges of E which start and end at vertices in V':

$$E' = \{(v_i, v_j) \in E \text{ for all } v_i, v_j \in V'\}.$$

A subgraph G' of G is an *undirected clique* if it is an undirected complete graph. We can now give a coordinate

independent definition of decentralized systems:

Definition 9 (Decentralized system). *A system of type of Equation* (8) *is* centralized *if its associated information flow is an undirected complete graph. Otherwise, it is decentralized.*

We now address the fact that some information loops are trivial, as described in the beginning of this section. A *path* of length k in a graph $G = (V, E)$ is an ordered list of vertices $v_1, \dots, v_k$, without repetitions except possibly for v_1 and v_k, such that $(v_i, v_{i+1}) \in E$ for $i = 1 \dots, k-1$. A path is *closed* or a *loop* if $v_1 = v_k$.

Definition 10 (Information loop). *A nontrivial information loop in a decentralized system with information flow graph $G = (V, E)$ is a closed path $(v_1, \dots, v_k)$ such that the subgraph G' generated by $V' = \{v_1, \dots, v_k\}$ is not an undirected clique.*

This definition takes into account the fact that when a graph is fully connected, even though loops will exist, their presence has no effect on the dynamics of the system.

The definition of information loop points towards the use of techniques from homological algebra to handle the information flow graph. We define here a combinatorial object, called simplicial complex, which allows us to make a connection between the structure of decentralized systems and algebraic topology.

A k-simplex is determined by k+1 vertices; we use the usual notation $[x_1, x_2, \dots, x_{k+1}]$ for the k-simplex with vertices $x_1, \dots, x_{k+1}$. A k-simplex has k+1 facets which are (k-1)-simplices, they are given by $[x_2, \dots, x_k], [x_1, x_3, \dots, x_k], \dots, [x_1, \dots, x_{k-1}]$.

Definition 11. *An* abstract simplicial complex $\mathcal{S}$ *is a combinatorial object consisting of a set of simplices such that any facet of a simplex $s \in \mathcal{S}$ is also in $\mathcal{S}$. The* k-skeleton *of a simplicial complex is the set of simplices of dimension k or less.*

We have the following definition:

Definition 12 (Information Flow Complex). *Consider the decentralized control system*

$$\dot{x} = \sum_{i=1}^{n} \sum_{j=1}^{n_i} u_{ij}(\delta_i(\mu); h_i(x)) g_{ij}(x)$$

where all the functions and vector fields involved are smooth. We assign to this system the simplicial complex with n vertices $h_1, h_2, \dots, h_n$ and facets given according to the following:

1) *There is an edge between x_i and x_j if*

$$g_{jk}(x) \cdot h_i(x) \neq 0 \text{ for any } k = 1 \dots n_j$$

2) *There is a k-simplex with vertices $x_1, \dots, x_k$ if*

$$g_{jk} \cdot h_i \neq 0 \text{ for any } g_{jk} \in \{g_{j1}, \dots, g_{jn_j}\}_{LA}, \quad \text{for all } i, j = 1..k$$

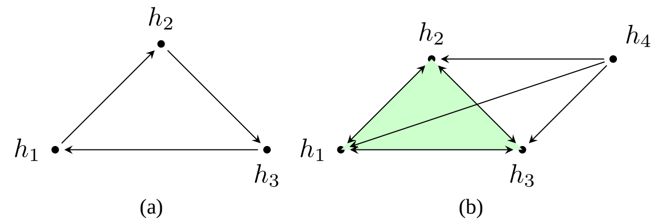


Fig. 5: In (a), we represent the information flow complex of system (5), which exhibits a nontrivial information loop. In (b), we represent the information flow complex of system (6); the shaded region depicts a 2-simplex in the complex. The information loop between $1, 2$ and 3 is trivial in this case.

The simplicial complex defined above is sometimes called a *Whitney complex* or *flag complex* in the literature. Due to space constraints, and the amount of background necessary to analyze such objects any further, most notably via their cohomology groups, we leave the study of the information flow complex to future work.

Example 11. *Consider the system*

$$\begin{cases} \dot{x}_1 &= u_1(x_1, x_2) \\ \dot{x}_2 &= u_2(x_2, x_3) \\ \dot{x}_3 &= u_3(x_3, x_1) \end{cases} \tag{5}$$

Using the notation introduced above, we have

$$h_1(x) = (x_1, x_2), h_2(x) = (x_2, x_3), h_3(x) = (x_3, x_1)$$

and $g_i(x) = e_i$. We thus have

$$\begin{aligned} g_1 \cdot h_2 &= [0, 0]; & g_1 \cdot h_3 &= [0, 1] \\ g_2 \cdot h_1 &= [0, 1]; & g_2 \cdot h_3 &= [0, 0] \\ g_3 \cdot h_1 &= [0, 0]; & g_3 \cdot h_2 &= [0, 1] \end{aligned}$$

We thus associate the information complex

$$S = \{h_1, h_2, h_3, [h_1, h_2], [h_2, h_3], [h_3, h_1]\}.$$

Consider the system

$$\begin{cases} \dot{x}_1 &= u_1(x_1, x_2, x_3) \\ \dot{x}_2 &= u_2(x_1, x_2, x_3) \\ \dot{x}_3 &= u_3(x_1, x_2, x_3, x_4) \\ \dot{x}_4 &= u_4(x_1, x_2, x_4) \end{cases} \tag{6}$$

Using the same approach as above, we find that the information complex of this system is (see Figure 5b)

$$S = \{h_1, h_2, h_3, h_4, [h_1, h_2], [h_2, h_3], [h_3, h_1], [h_1, h_4], [h_2, h_4], [h_3, h_4], [h_1, h_2, h_3]\}.$$

References

[1] N. R. Sandell, P. Varaiya, M. Athans, and M. G. Safonov, "A survey of decentralized control methods for large scale systems," *IEEE T. Automatic Control*, vol. 23, no. 2, pp. 108–128, 1978.

[2] M. Rotkowitz and S. Lall, "A characterization of convex problems in decentralized control," *IEEE T. Automatic Control*, vol. 51, no. 2, pp. 274–286, 2006.

[3] P. Shah and P. A. Parrilo, "A poset framework to model decentralized control problems," in *Proc. of the 48th IEEE Conference on Decision and Control*, vol. 48, 2009.

[4] P. Krishnamurthy, F. Khorrami, and D. A. Schoenwald, "Decentralized inventory control for large-scale reverse supply chains: A computationally tractable approach," *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, vol. 38, no. 4, pp. 551–561, 2008.

[5] J. A. Fax and R. M. Murray, "Information flow and cooperative control of vehicle formations," *IEEE T. Automatic Control*, vol. 49, no. 9, pp. 1465–1476, 2004.

[6] C. Yu, B. Anderson, S. Dasgupta, and B. Fidan, "Control of minimally persistent formations in the plane," *SIAM Journal on Control and Optimization*, vol. 48, no. 1, pp. 206–233, 2009.

[7] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithms and theory," *IEEE T. Automatic Control*, vol. 51, no. 3, pp. 401–420, 2006.

[8] S. Seuken and S. Zilberstein, "Formal models and algorithms for decentralized decision making under uncertainty," *Autonomous Agents and Multi-Agent Systems*, vol. 17, no. 2, pp. 190–250, 2008.

[9] J. K. Hedrick, M. Tomizuka, and P. Varaiya, "Control issues in automated highway systems," *IEEE Control Systems Magazine*, vol. 14, pp. 21–32, 1994.

[10] P. Barooah, P. G. Mehta, and J. P. Hespanha, "Mistuning-based control design to improve closed-loop stability margin of vehicular platoons," *IEEE Trans. on Automat. Contr.*, vol. 54, no. 9, pp. 2100–2113, 2009.

[11] M.-A. Belabbas, "Mathematical aspects of decentralized control of formations in the plane." in *Proceedings of the 9th International Conference on Control and Automation*, 2011.

[12] ——, "Geometric properties of formation control," in *Submitted , available on my website*, 2011.

[13] L. Krick, M. E. Broucke, and B. Francis, "Stabilization of infinitesimally rigid formations of multi-robot networks," *International Journal of Control*, 2008.

[14] R. Connelly, "Generic global rigidity," *Discrete Comput. Geom.*, vol. 33, pp. 549–563, 2005.

[15] G. Birkhoff, *Lattice theory*. American Mathematical Society, 1948.

[16] F. Warner, *Foundations of Differentiable Manifolds and Lie Groups*. Springer, 1983.

[17] R. W. Brockett, "Asymptotic stability and feedback stabilization," in *Differential Geometric Control Theory*, R. S. M. R. W. Brockett and H. J. Sussmann, Eds. Boston: Birkhauser, 1983, pp. 181–191.

[18] V. D. Silva and R. Ghrist, "Homological sensor networks," *Notices of the American Mathematical Society*, vol. 54, p. 2007, 2007.

[19] H. Edelsbrunner, "Biological applications of computational topology," in *Handbook of Discrete and Computational Geometry*, J. E. Goodman and J. O'Rourke, Eds. Boca Raton: CRC Press,, 2004, pp. 1395–1412.

[20] R. Ghrist, "Barcodes: The persistent topology of data," Tech. Rep., 2007.